

Concept URI Strategy for the NL Public Sector

DRAFT version – Q1 2013

Hans Overbeek

Thijs Brentjens

Content

Starting points
The 'quest'
Conclusion
Details



Starting points

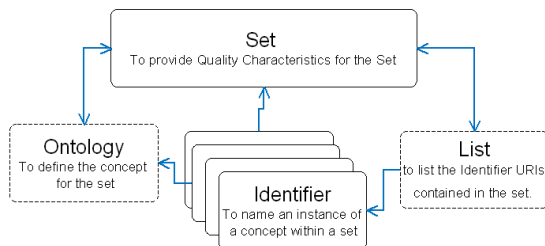
Directive, Article 8(2)(a)



"The implementing rules shall address a common framework for the unique identification of spatial objects, to which identifiers under national systems can be mapped"

ISA: D7.1.3 - Study on persistent URIs, with identification of best practices and recommendations on the topic for the MSs and the EC

Designing URI Sets for the UK Public Sector.



Follow the pattern
e.g. `http://(domain)/(type)/(concept)/(reference)`

Re-use existing identifiers
e.g. `http://education.data.gov.uk/id/school/123456`

Link multiple representations
e.g. `http://data.example.org/doc/foo/bar.html`
e.g. `http://data.example.org/doc/foo/bar.rdf`

Implement 303 redirects for real-world objects
e.g. `http://www.example.com/id/alice_brown`

Use a dedicated service
i.e. independent of the data originator

10
rules
for **persistent**
URIs

Avoid stating ownership
e.g. `http://education.data.gov.uk/ministryXeducation/id/school/123456`

Avoid version numbers
e.g. `http://education.data.gov.uk/doc/school/123456`

Avoid using auto-increment
e.g. `http://education.data.gov.uk/id/school/123456`
e.g. `http://education.data.gov.uk/id/school/123457`

Avoid query strings
e.g. `http://education.data.gov.uk/doc/school?id=123456`

Avoid file extensions
`http://education.data.gov.uk/doc/schools/123456.c`

The Quest



Functions of Linked Data



A. Standard (ex: SKOS, DC, FOAF, ECLI, OWMS)

- Semantic model (ontology) for a domain or sector
- Vocabulary for the model

B. Register (ex: roads, schools)

- Administration of 'Things'
- Mints URI's for things, to be re-used
- Uses vocabulary of the standards

C. Application (ex: weather, trafic, population)

- Provides data about the things in the registers
- Uses URI's from register and vocabulary of the standards

Goal of URI-strategy

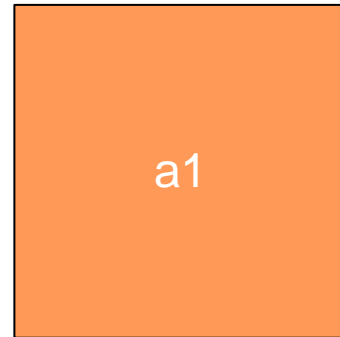


Standard

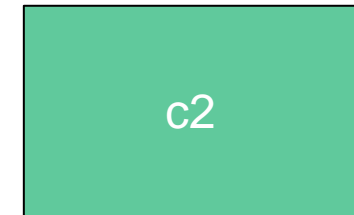
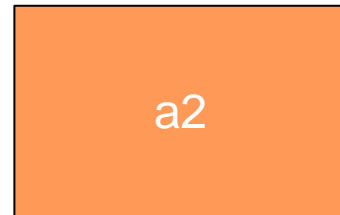
Register

Application

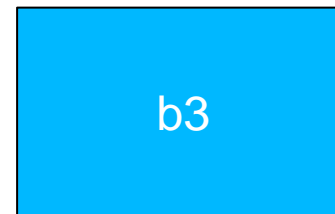
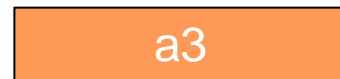
Model



**Reference-
objects**



Data



Re-use of Ontology Terms



Standard

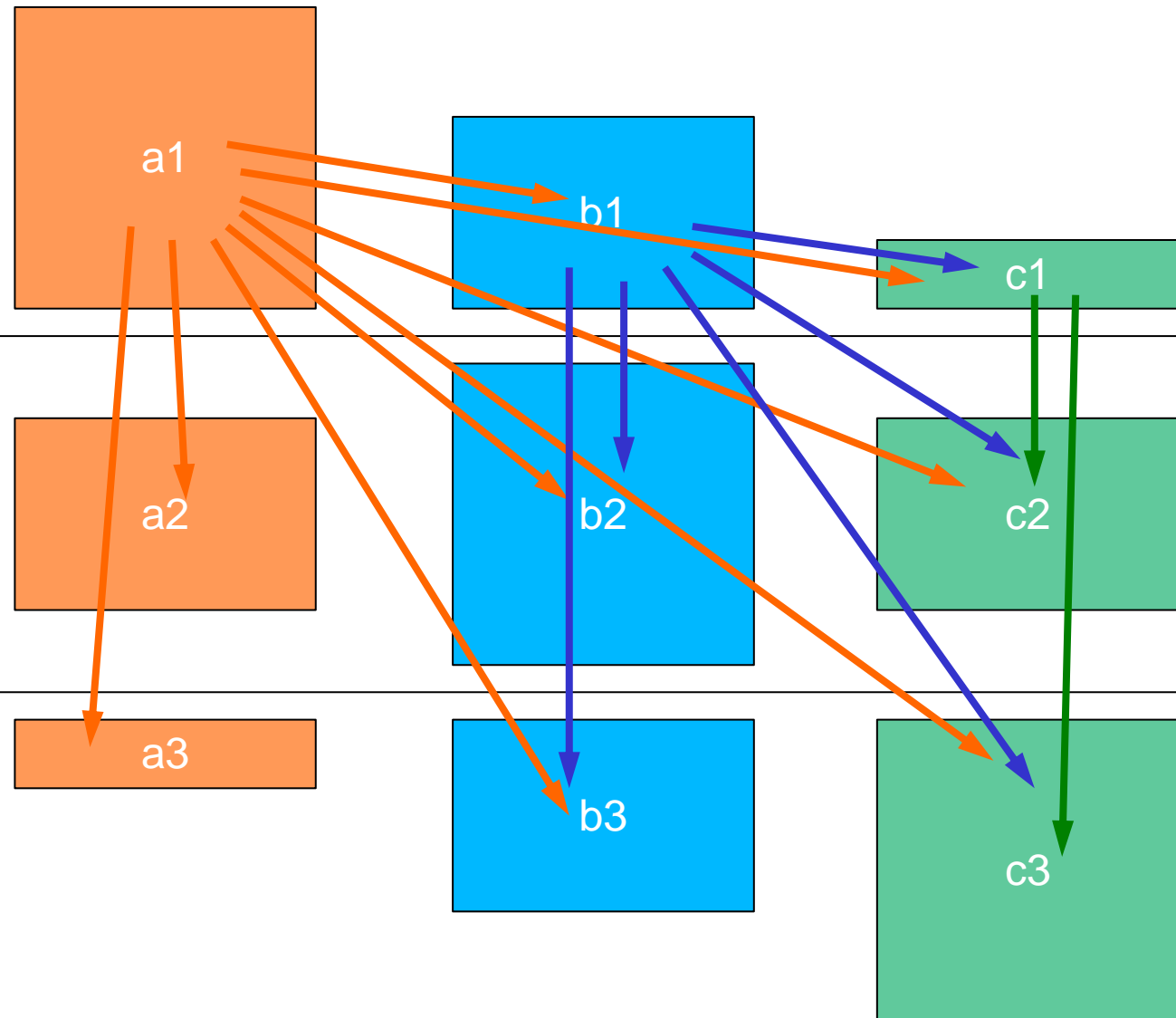
Register

Application

Model

**Reference-
objects**

Data



Re-use of Reference Objects

Insight



Standard

Register

Application

Model

a1

b1

c1

Reference-objects

a2

b2

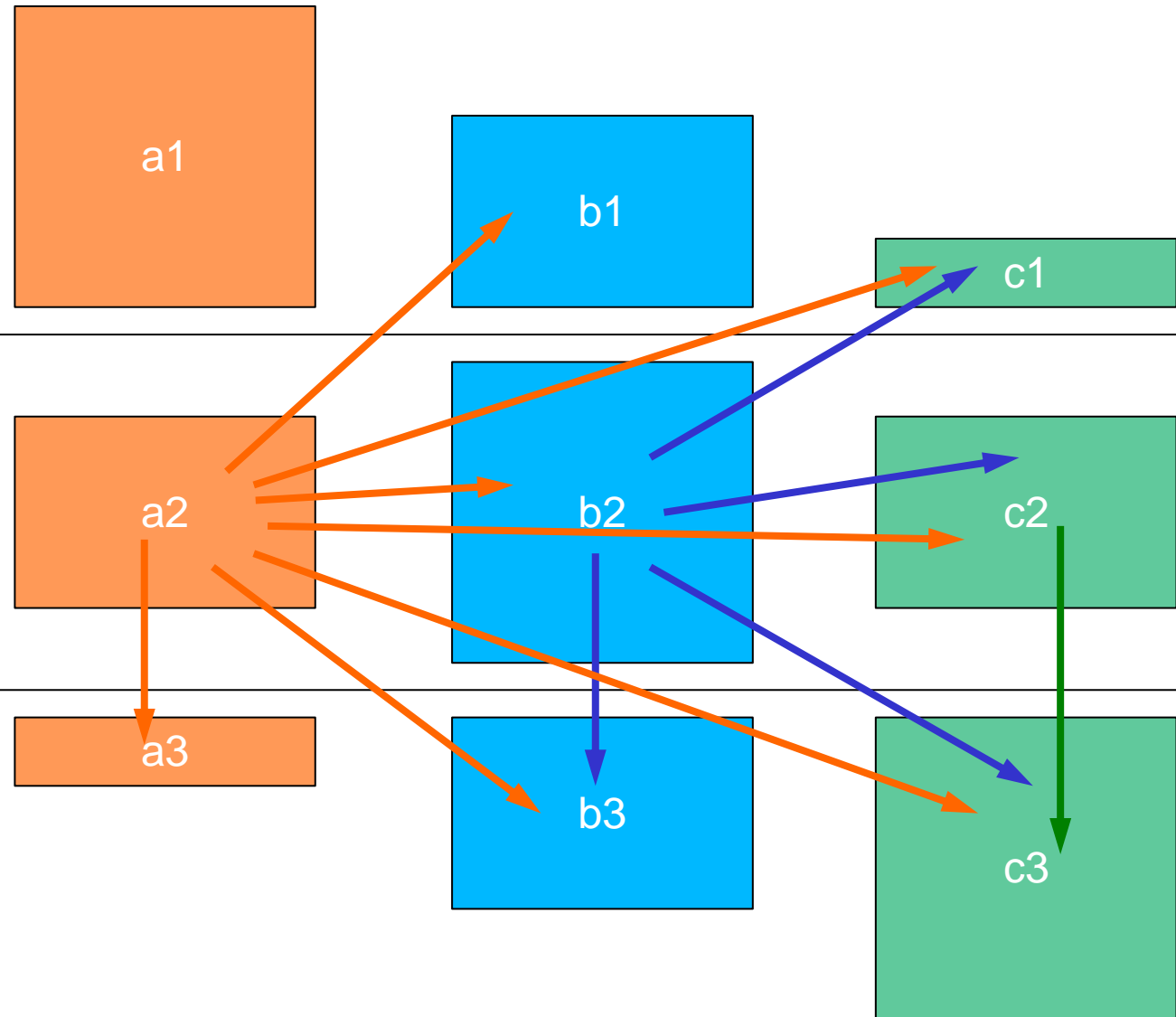
c2

a3

b3

c3

Data

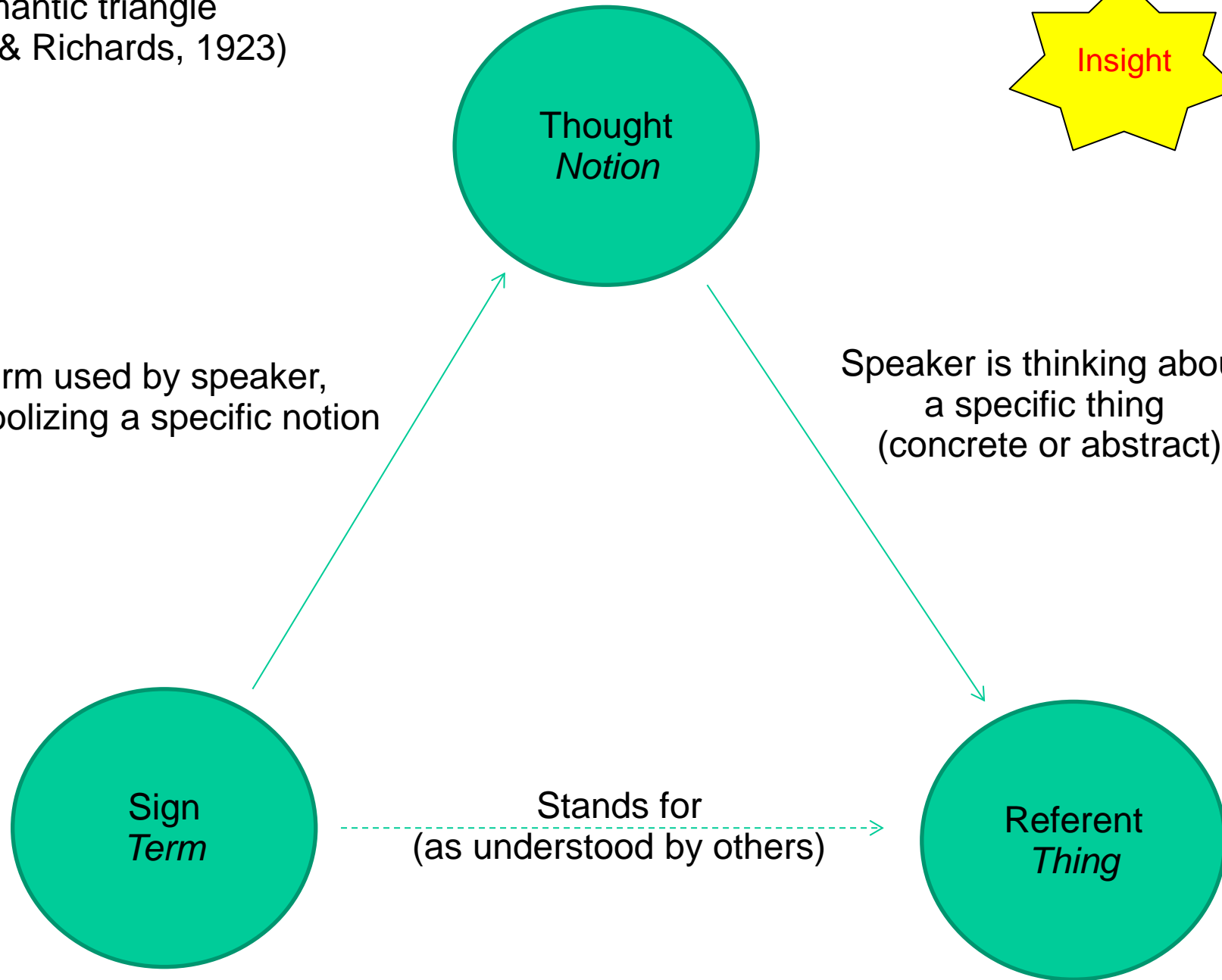


Semantic triangle
(Ogden & Richards, 1923)



Term used by speaker,
symbolizing a specific notion

Speaker is thinking about
a specific thing
(concrete or abstract)

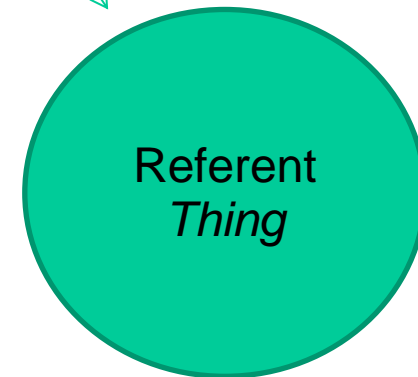
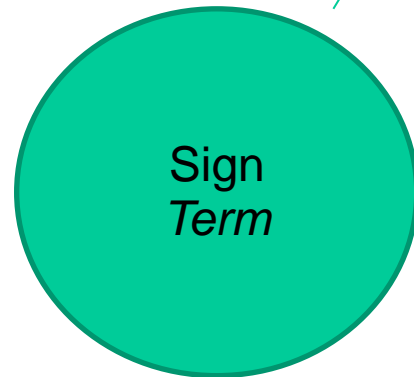


Semantic triangle
(Ogden & Richards, 1923)



Term used by speaker,
symbolizing a specific notion

Speaker is thinking about
a specific thing
(concrete or abstract)



Stands for
(as understood by others)

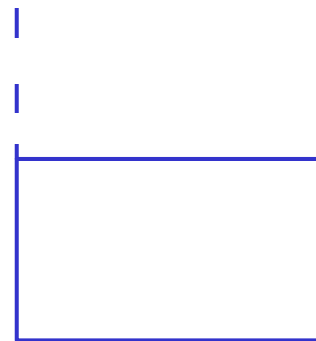
URI from a register



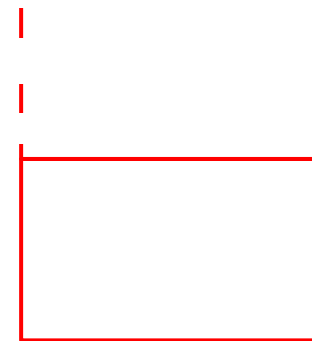
No Register?
No Identifier!

One URI to Unite Them All?

Insight



BAG (contour)



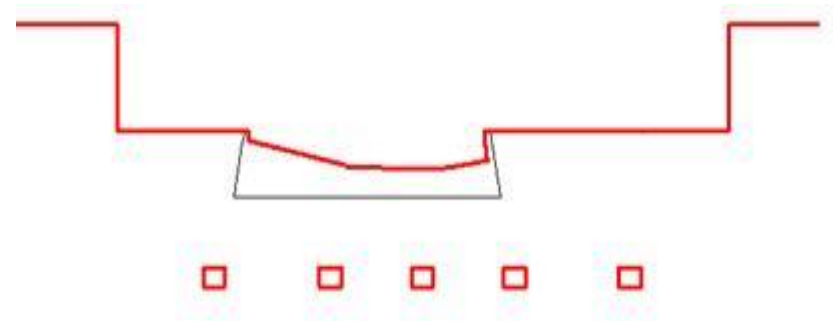
BGT (footprint)

One URI to Unite Them All?

Insight

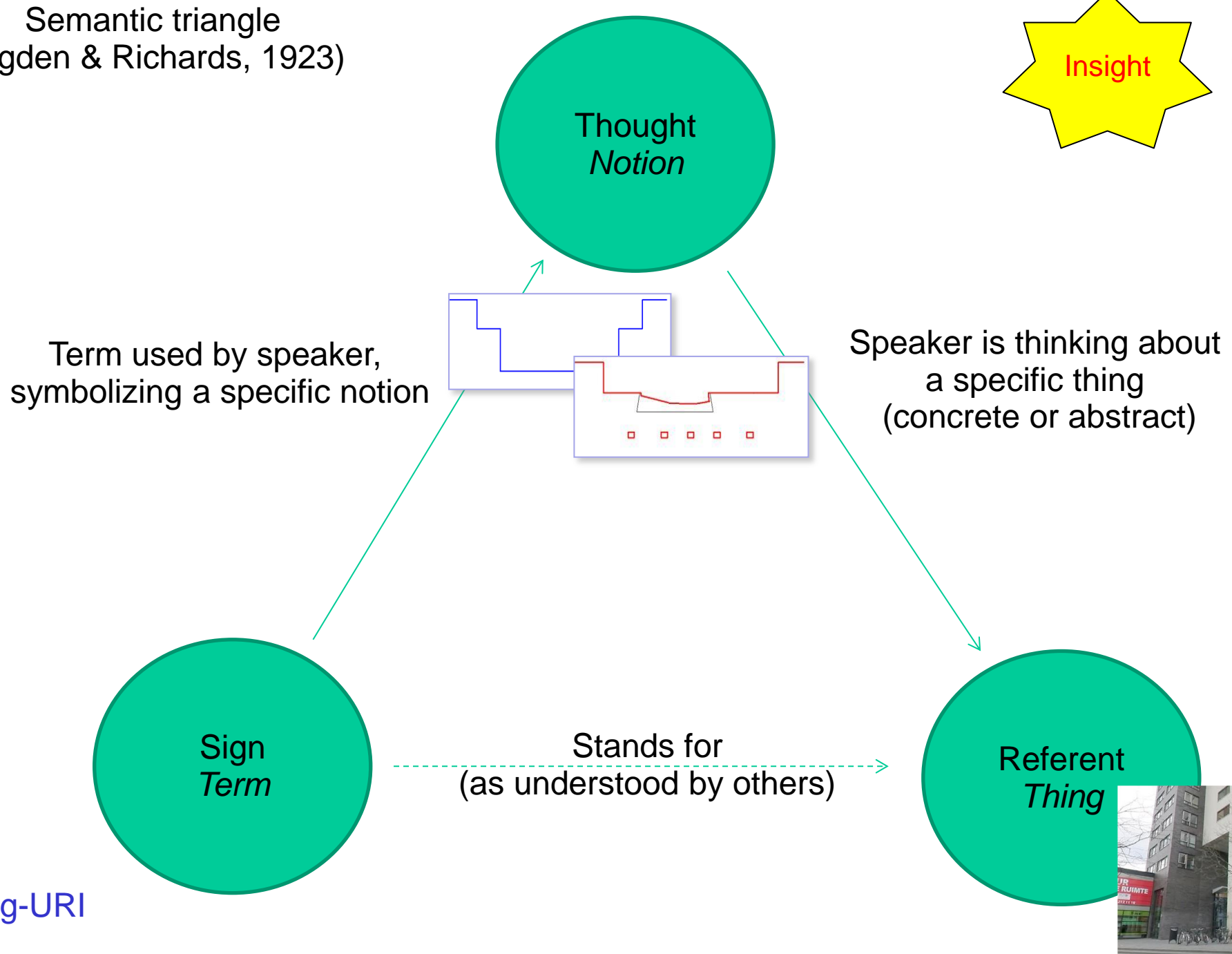


BAG (contour)



BGT (footprint)

Semantic triangle
(Ogden & Richards, 1923)



bag-URI

Bgt-URI

http://en.wikipedia.org/wiki/Triangle_of_reference

Conclusion

- KISS (human readable, short)
- Trustworthy
- Intuitive
- Persistent

<http://{domain}/{type}/{concept}/{reference}>

<http://{domain}/{type}/{concept}/{reference}>

{domain} identifies the register in a persistent way

So ideally: {register}.{top-domain}/

or: {sector}.{top-domain}/{register} ?

or {yourdomain}/{yourpath} ?

{top-domain}:

- 'data.gov.nl': recognisable, trustworthy, persistent
- {register}.'data.pilod.nl' voor de PiLOD

Questions

Which part DNS-adressable?

{register}: No register, no identifier!

<http://{domain}/{type}/{concept}/{reference}>

{domain} = {sector}.data.pilod.nl/{register}

What infrastructure is needed?

- A register of sectors (which sectors?)
 - Sector name
 - Sector mnemonic
 - Sector resolver
 - Sector administrator
- How will apps use the resolvers and how frequently?
- Complications with {register} at the end of {domain}?



Questions

<http://{domain}/{type}/{concept}/{reference}>

{type}

- 'id': identifier of real life object *in a register*
- 'doc': documentation about the real life object by this register
- 'def':

Wait, let me...

<http://{domain}/{type}#{concept}>

{type}

- ‘def’: definition of a term in an ontology
- Hash-URI
- URI of model: <http://{domain}/def>

UK-strategy recommends slash URI's for vocabulary terms. Why?

Sometimes no clear distinction between model and content

Questions

<http://{domain}/{type}/{concept}/{reference}>

{concept}

- The {concept} is 'just a string' NOT the formal classification of the identified resource. It has no semantic meaning, just comes in handy for the human reader and helps to make unique references
- Singular
- Avoid all other characters than a-z, A-Z, 0-9
- CamelCase or lowercase?

Questi
ons

`http://{domain}/def#{concept}`

To identify concepts:

- Classes, properties and skos:concepts
- Hash-URI always returns complete ontology
`http://{domain}/def`
- Use UpperCamelCase for classes and concepts, lowerCamelCase for properties

`http://{domain}/{type}/{concept}/{reference}`

`{reference}`

Key of the resource within the register. URI-strategy leaves lot of freedom to the registrar: many different requirements

Some recommendations:

- Try to avoid special characters
- Use W3CDTF for versions (yyyy-mm-dd)
- Avoid information bearing keys

<http://{domain}/{type}/{concept}/{reference}>

Object-identifiers

- <http://bag.data.pilod.nl/id/pand/12345>
- <http://bgt.data.pilod.nl/id/pand/86420>

Documentatie

- <http://bag.data.pilod.nl/doc/pand/12345>

Lijst

- <http://bgt.data.pilod.nl/id/pand>

Concept

- <http://bgt.data.pilod.nl/def#pand>

Model

- <http://bgt.data.pilod.nl/def#pand>